Math 2270 - Lecture 40 : Markov Matrices

Dylan Zwick

Fall 2012

This lecture covers section *section 8.3* of the textbook.

Today we're going to talk about a very special set of matrices called *Markov matrices*. These are matrices where every entry $a_{ij} > 0$. These matrices come up all the time in statistics.

The assigned problems for this section are

Section 8.3 - 1, 2, 3, 9, 10

Markov Matrices

A Markov matrix is a type of matrix that comes up in the context of something called a *Markov chain* in probability theory. A Markov matrix is a square matrix with all nonnegative entries, and where the sum of the entries down any column is 1. If the entries are all positive, it's a *positive* Markov matrix.

The most important facts about a positive Markov matrix are:

- $\lambda = 1$ is an eigenvalue.
- The eigenvector associated with $\lambda = 1$ can be chosen to be strictly positive.
- All other eigenvalues have magnitude less than 1.

Before diving into these, let's take a look at some basic properties of Markov matrices. Suppose $\mathbf{u}_0$ is a positive vector (all components are positive) whose components add to 1. This is true if and only if $\begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \mathbf{u}_0 = 1$. If A is a positive Markov matrix, then the product $\mathbf{u}_1 = A\mathbf{u}_0$ will also be a positive vector with components adding to 1. This is because if A is a matrix all of whose columns add to 1, then the product of the matrix with the row vector $\begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}$ will be the row vector consisting of all 1s. From this we get:

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} A\mathbf{u}_0 = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \mathbf{u}_0 = 1.$$

So, we see $\mathbf{u}_1 = A\mathbf{u}_0$ is a positive vector with components adding to 1, and by induction $A^k\mathbf{u}_0$ is a positive vector with components adding to 1 for every k.

But, what is this vector? Does is oscillate, or does $A^k\mathbf{u}_0$ approach some steady state vector $\mathbf{u}_\infty$. If it does, what is this vector?

It turns out that if the Markov matrix is positive, it approaches a steady state, and this steady state is given by the eigenvector associated with $\lambda = 1$. The reason for this is that we can write $\mathbf{u}_0$ as a linear combination of our eigenvectors:

$$\mathbf{u}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n.$$

The product $A\mathbf{u}_0$ will be

$$A\mathbf{u}_0 = c_1A\mathbf{v}_1 + c_2A\mathbf{v}_2 + \cdots + c_nA\mathbf{v}_n = c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \cdots + c_n\lambda_n\mathbf{v}_n$$

and in general

$$A^k\mathbf{u}_0 = c_1\lambda_1^k\mathbf{v}_1 + \cdots + c_n\lambda_n^k\mathbf{v}_n.$$

We know $\lambda_1 = 1$ and $|\lambda_i| < 1$ for $i > 1$, and so as $k \to \infty$ the vector $\mathbf{u}_0$ will approach $c_1\mathbf{v}_1$, which will be the steady state solution. Note this assumes $c_1 \neq 0$.

So, why is $\lambda = 1$ an eigenvalue of a Markov matrix? Because every column of $A - I$ adds to $1 - 1 = 0$. So, the rows of $A - I$ add up to the zero row, and those rows are linearly dependent, so $A - I$ is singular. So, λ is an eigenvalue of A.

The reason no eigenvalue of A has $|\lambda| > 1$ is that the powers of A^k would grow. But, every A^k must also be a Markov matrix, and so it can't get large.[1]

That we can find a positive eigenvector for $\lambda = 1$ follows from the Perron-Frobenius theorem. An awful and not really correct proof of this theorem can be found in the textbook.

Example - What is the steady state for the Markov matrix

$$A = \begin{pmatrix} .80 & .05 \\ .20 & .95 \end{pmatrix}$$

$$\begin{vmatrix} \frac{4}{5} - \lambda & \frac{1}{20} \\ \frac{1}{5} & \frac{19}{20} - \lambda \end{vmatrix} = \left(\frac{4}{5} - \lambda\right)\left(\frac{19}{20} - \lambda\right) - \frac{1}{100}$$

$$= \lambda^2 - \frac{35}{20}\lambda + \frac{76}{100}\cancel{\lambda} - \frac{1}{100} = \lambda^2 - \frac{7}{4}\lambda + \frac{3}{4}$$

$$= (\lambda - 1)\left(\lambda - \frac{3}{4}\right)$$

$$\begin{pmatrix} -\frac{1}{5} & \frac{1}{20} \\ \frac{1}{5} & -\frac{1}{20} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ 4 \end{pmatrix} \qquad \begin{pmatrix} \frac{1}{5} \\ \frac{4}{5} \end{pmatrix}$$

$$\begin{pmatrix} \frac{4}{5} & \frac{1}{20} \\ \frac{1}{5} & \frac{19}{20} \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

[1] This is not a formal proof, but it's the basic idea.